

Figure 1: cDNA sequence (and encoded amino acid sequence) of Cynomologous monkey Cathepsin S

```

1  AGTTGAACTGAAATCTCCCTGCTGCTGCTTTGAAATCTTAGAAGAGAGCCCATCAATTCA
61  AGGATTCTTACTGTAGGAGCACCTGCTGGTTCTATCACAATGAAACAGCTGGTTTGTGTG
      M K Q L V C V
      ◆
121  CTCTTGGTGTGCTCCTCTGCAGTGGCGCAGTTGCATAAAGATCCTACCCTGGATCATCAC
      L L V C S S A V A Q L H K D P T L D H H
181  TGGCATCTCTGGAAGAAAACCTATGGCAAACAATAAAGGAAAAGAATGAAGAAGCAGTA
      W H L W K K T Y G K Q Y K E K N E E A V
241  CGACGTCTCATCTGGGAAAAGAATCTAAAGTTTGTGATGCTTCATAACCTGGAGCATTCA
      R R L I W E K N L K F V M L H N L E H S
301  ATGGGAATGCACTCATATGATCTGGGCATGAACCACCTGGGAGACATGACCAGTGAAGAA
      M G M H S Y D L G M N H L G D M T S E E
361  GTGATGTCCTTGATGAGTTCCCTGAGAGTTCCAGCCAGTGGCAGAGAAATATCACATAT
      V M S L M S S L R V P S Q W Q R N I T Y
      ↓
421  AAGTCAAACGCTAATCAGATATTGCCGGATTCTGTGGACTGGAGAGAGAAAGGGTGTGTT
      K S N A N Q I L P D S V D W R E K G C V
481  ACTGAAGTGAATATCAAGGTTCTTGTGGTGTCTGCTGGGCTTTCAGTGTCTGGGGGCC
      T E V K Y Q G S C G A C W A F S A V G A
541  CTGGAAGCACAGCTGAAGCTGAAAACAGGAAGCTGGTGTCTCTCAGTGGCCAGAACCTG
      L E A Q L K L K T G K L V S L S A Q N L
601  GTGGATTGCTCAACTGAAAATATGGAACAAAGGCTGCAATGGTGGCTTCATGACAAGG
      V D C S T E K Y G N K G C N G G F M T R
661  GCTTTCCAGTACATCATTGATAACAACGGCATCGACTCAGACGCTTCCTATCCCTACAAA
      A F Q Y I I D N N G I D S D A S Y P Y K
721  GCCACGGATCAGAAGTGTCAATATGACTCAAATATCGTGTGCCACATGTTCAAAGTAT
      A T D Q K C Q Y D S K Y R A A T C S K Y
781  ACTGAACTTCCTTATGGCAGAGAAGATGTCCTGAAAGAAGTTGTGGCCAATAAAGGCCCA
      T E L P Y G R E D V L K E V V A N K G P
841  GTGTCTGTTGGTGTGGATGCGAGTCATCCTTCTTCTCTCTACAGAAGTGGTGTCTAC
      V S V G V D A S H P S F F L Y R S G V Y
901  TATGAACCATCCTGTACTCAGAATGTGAATCATGGTGTACTTGTGGTGGCTATGGTGT
      Y E P S C T Q N V N H G V L V V G Y G V
961  CTTAACGGGAAAGAATACTGGCTTGTGAAAAACAGCTGGGGCCGCAACTTTGGTGAAGAA
      L N G K E Y W L V K N S W G R N F G E E
1021  GGATATATTCCGATGGCAAGAAATAAAGGAAATCATTGTGGGATTGCTAGTTCCCTCT
      G Y I R M A R N K G N H C G I A S F P S
1081  TACCCAGAAATCTAGAGAGGATCTCTTCTTTTATAACAAATCAAGAAATATGAAGCAC
      Y P E I *
1141  TTTCTCTTAACCTAATTTTCTCTGCTGTATCCAGAAGAAATAATTGTGTACCGATTAATG
1201  TGTATTTACTGTACTAATTAATAAATATAGTTTGAGGCCGGGCACCGTGGCTCACGCCTG
1261  TAATCCCACTACTTTGGGAGGCCAAGGCAGGCATATCAACTTGAGGCCAGGAGTTAAAGA
1321  CAGGCCTGGCTAACATGGTGAACCCCGTCTCTACTAAAAATACAAAACATTAGCGGAGC
1381  GTAATGGTGCATGCCGTGAATCCCACTACTTGGGAGGCTGAAGCACAAGATTCTTTGAA
1441  CCCAAGAGTTGAGGCTGTGGTGAAGTGAAGACCACCACTGTACTCCAGCCTGGACAAC
1501  AGAGTGGAGACTCTGTTTCAAAAAACAGAAAAGATAATATAGTTTGATTCTTCGTTT
1561  TAAAAATTTGCAAACTCAGGAAAAAGTTTGTAAAGTAAATTAGTAGTGTACTATAGATAT
1621  AACTGTATAAAAAATTGTTCAACCTAAAACAATCTGTATTGTTTATTTTATTTTATA
1681  CTCTTTGTCTTTTAAAGACCCTGATAGCCTTTTGTAACTTGATGGCTTAAAGTACTTA
1741  ATAAATCTGCCATTTCAAATTTCAATTCAAAAA

```

SEQ ID NO:2

SEQ ID NO:1

◆ = cleavage site for pre-sequence

↓ = cleavage site for pro-sequence

\* = stop codon

FIGURE 1

Figure 2: Alignment of cathepsin S amino acid sequences from Cynomologous monkey (SEQ ID NO:2, as shown in Fig. 1), human (Genbank accession number BC002642.2), monkey sml cells (*Saimiri boliviensis*), dog (*Canis familiaris*; Genbank accession number AY156692), and mouse (*Mus musculus*; Genbank gi3850787), followed by a consensus sequence.

“cyno cats preprotein” = Cynomologous monkey = SEQ ID NO:2

“smlmonkey cats” = monkey sml cells (*Saimiri boliviensis*) = SEQ ID NO:3

“hcats prepropro.txt” = human (*Homo sapiens*) = SEQ ID NO:4

“dog cats” = dog (*Canis familiaris*) = SEQ ID NO:5

“mouse cats prepro” = mouse (*Mus musculus*) = SEQ ID NO:6

(consensus sequence = SEQ ID NO:7)

FIGURE 2A

## ClustalW Formatted Alignments

	10	20	30	40
<i>cyno cats preprotein</i>	M K Q L V C V L L V C S S A V A Q L H K D P T L D H H W H L W K K T Y G K Q Y K			
<i>smmonkey cats</i>	M K Q L V C V L F V C S S A V T Q L H K D P T L D H H W N L W K K T Y G K Q Y K			
<i>hcats prepropro.txt</i>	M K R L V C V L L V C S S A V A Q L H K D P T L D H H W H L W K K T Y G K Q Y K			
<i>dog cats</i>	M K W L V G L L P L C S Y A V A Q V H K D P T L D H H W N L W K K T Y S K Q Y K			
<i>mouse catsprepro</i>	M A V L D A P G V L C G N G A T A E R . . P T L D H H W D L W K K T H E K E Y K			
	M K . L V C V L . V C S S A V A Q L H K D P T L D H H W . L W K K T Y G K Q Y K			
	50	60	70	80
<i>cyno cats preprotein</i>	E K N E E A V R R L I W E K N L K F V M L H N L E H S M G M H S Y D L G M N H L			
<i>smmonkey cats</i>	E K N E E A V R R L I W E K N L K F V M L H N L E H S M G M H S Y D L G M N H L			
<i>hcats prepropro.txt</i>	E K N E E A V R R L I W E K N L K F V M L H N L E H S M G M H S Y D L G M N H L			
<i>dog cats</i>	E E N E E V A R R L I W E K N L K F V M L H N L E H S M G M H S Y D L G M N H L			
<i>mouse catsprepro</i>	D K N E E E V R R L I W E K N L K F I M I H N L E Y S M G M H T Y Q V G M N D M			
	E K N E E A V R R L I W E K N L K F V M L H N L E H S M G M H S Y D L G M N H L			
	90	100	110	120
<i>cyno cats preprotein</i>	G D M T S E E V M S L M S S L R V P S Q W Q R N I T Y K S N A N Q I L P D S V D			
<i>smmonkey cats</i>	G D M T S E E V M S L M S S L R V P N Q W Q R N I T Y K S N P N Q M L P D S V D			
<i>hcats prepropro.txt</i>	G D M T S E E V M S L M S S L R V P S Q W Q R N I T Y K S N P N W I L P D S V D			
<i>dog cats</i>	G D M T G E E V I S L M G S L R V P S Q W Q R N V T Y R S N S N Q K L P D S V D			
<i>mouse catsprepro</i>	G D M T I N E E I L C R M G A L R I I P R Q S P K T V T F R S Y S N R T L P D T V D			
	G D M T S E E V M S L M S S L R V P S Q W Q R N I T Y K S N . N Q . L P D S V D			
	130	140	150	160
<i>cyno cats preprotein</i>	W R E K G C V T E V K Y Q G S C G A C W A F S A V G A L E A Q L K L K T G K L V			
<i>smmonkey cats</i>	W R E K G C V T E V K Y Q G S C G A C W A F S A V G A L E A Q L K L K T G K L V			
<i>hcats prepropro.txt</i>	W R E K G C V T E V K Y Q G S C G A C W A F S A V G A L E A Q L K L K T G K L V			
<i>dog cats</i>	W R E K G C V T E V K Y Q G S C G A C W A F S A V G A L E A Q L K L K T G K L V			
<i>mouse catsprepro</i>	W R E K G C V T E V K Y Q G S C G A C W A F S A V G A L E G Q L K L K T G K L I			
	W R E K G C V T E V K Y Q G S C G A C W A F S A V G A L E A Q L K L K T G K L V			
	170	180	190	200
<i>cyno cats preprotein</i>	S L S A Q N L V D C S T E K Y G N K G C N G G F M T R I A F Q Y I I D N N G I D			
<i>smmonkey cats</i>	S L S A Q N L V D C S T . E K Y G N K G C N G G F M T E A F Q Y I I D N K G I D			
<i>hcats prepropro.txt</i>	S L S A Q N L V D C S T E K Y G N K G C N G G F M T T A F Q Y I I D N K G I D			
<i>dog cats</i>	S L S A Q N L V D C S T . E K Y G N K G C N G G F M T T A F Q Y I I D N N G I D			
<i>mouse catsprepro</i>	S L S A Q N L V D C S N E E K Y G N K G C G G G Y M T E A F Q Y I I D N G I E			
	S L S A Q N L V D C S T E K Y G N K G C N G G F M T . A F Q Y I I D N . G I D			
	210	220	230	240
<i>cyno cats preprotein</i>	S D A S Y P Y K A T D Q K C Q Y D S K Y R A A T C S K Y T E L P Y G R E D V L K			
<i>smmonkey cats</i>	S E A S Y P Y K A T D Q K C Q Y D S K Y R A A T C S K Y T E L P Y G R E D V L K			
<i>hcats prepropro.txt</i>	S D A S Y P Y K A M D Q K C Q Y D S K Y R A A T C S K Y T E L P Y G R E D V L K			
<i>dog cats</i>	S E A S Y P Y K A M N G K C R Y D S K K R A A T C S K Y T E L P F G S E D A L K			
<i>mouse catsprepro</i>	A D A S Y P Y K A M D E K C H Y N S K N R A A T C S R Y I Q L P F G D E D A L K			
	S D A S Y P Y K A M D Q K C Q Y D S K Y R A A T C S K Y T E L P Y G R E D V L K			
	250	260	270	280
<i>cyno cats preprotein</i>	E V V A N K G P V S V G V D A S H P S F F L Y R S G V Y Y E P S C T Q N V N H G			
<i>smmonkey cats</i>	E A V A N K G P V C V G V D A S H P S F F L Y R S G V Y Y D P A C T Q K V N H G			
<i>hcats prepropro.txt</i>	E A V A N K G P V S V G V D A R H P S F F L Y R S G V Y Y E P S C T Q N V N H G			
<i>dog cats</i>	E A V A N K G P V S V A I D A S H Y S F F L Y R S G V Y Y E P S C T Q N V N H G			
<i>mouse catsprepro</i>	E A V A T K G P V S V G I D A S H S S F F Y K S G V Y D D P S C T G N V N H G			
	E A V A N K G P V S V G V D A S H P S F F L Y R S G V Y Y E P S C T Q N V N H G			
	290	300	310	320
<i>cyno cats preprotein</i>	V L V V G Y G V L N G K E Y W L V K N S W G R N F G E E G Y I R M A R N K G N H			
<i>smmonkey cats</i>	V L V I G Y G D L N G K E Y W L V K N S W G S N F G E Q G Y I R M A R N K G N H			
<i>hcats prepropro.txt</i>	V L V V G Y G D L N G K E Y W L V K N S W G H N F G E E G Y I R M A R N K G N H			
<i>dog cats</i>	V L V V G Y G N L N G K D Y W L V K N S W G L N F G D Q G Y I R M A R N S G N H			
<i>mouse catsprepro</i>	V L V V G Y G T L D G K D Y W L V K N S W G L N F G D Q G Y I R M A R N N K N H			
	V L V V G Y G . L N G K E Y W L V K N S W G . N F G E Q G Y I R M A R N K G N H			
	330	340	350	360
<i>cyno cats preprotein</i>	C G I A S F P S Y P E I			
<i>smmonkey cats</i>	C G I A S Y P S Y P E I			
<i>hcats prepropro.txt</i>	C G I A S F P S Y P E I			
<i>dog cats</i>	C G I A S Y P S Y P E I			
<i>mouse catsprepro</i>	C G I A S Y C S Y P E I			
	C G I A S Y P S Y P E I			

FIGURE 2B

Figure 3: Purification of monkey Cathepsin S (Gelcode stained gel)

